

PatchMatch多視点ステレオの高精度化に関する研究

著者	伊藤 隆史
雑誌名	東北大学電通談話会記録
巻	90
号	1
ページ	230-231
発行年	2021-08-20
URL	http://hdl.handle.net/10097/00132894

修士学位論文要約（令和3年3月）

PatchMatch 多視点ステレオの高精度化に関する研究

伊藤 隆史

指導教員：青木 孝文

Accuracy Improvement of PatchMatch Multi-View Stereo

Takafumi ITO

Supervisor: Takafumi AOKI

Multi-View Stereo (MVS) is a technique to reconstruct accurate and dense 3D shapes from multi-view images. PatchMatch Multi-View Stereo (PM-MVS) is one of the MVS methods that can estimate high-accuracy depth maps and normal maps by initializing parameters with random values and propagating parameters based on photoconsistency of a set of multi-view images. On the other hand, its reconstruction accuracy is degraded due to various factors such as unstructured camera geometry and heterogeneous scene illuminations. Addressing the above problem, this paper proposes three methods to improve the reconstruction accuracy of PM-MVS. Through the experiments using the ETH3D dataset and self-made dataset including human ear images, we demonstrate the effectiveness of proposed methods.

1. はじめに

多視点ステレオ (Multi-View Stereo: MVS) は、物体を様々な位置や姿勢から撮影した複数枚の画像を入力として対象の 3 次元形状の復元を行う手法である。MVS 手法の 1 つである PatchMatch 多視点ステレオ (PatchMatch Multi-View Stereo: PM-MVS)¹⁾ は、乱数によるパラメータの初期化と多視点画像間の画像マッチングに基づくパラメータの伝播により、高速かつ高精度に復元を行うことが可能である。一方で、物体を撮影するカメラの位置や撮影環境が画像ごとに大きく異なる場合には、オクルージョンや画像の明度変化の影響が大きくなるため、高精度な復元を行うことができない。そこで、本論文では、PM-MVS を高精度化するために 3 つの高精度化手法を提案する²⁾。1 つ目の手法では、多視点画像間で計算されたマッチングスコア、視点間の幾何学的関係およびメッシュに基づいて、画素単位でマッチングに用いる視点を選択することで、復元精度を向上させる。2 つ目の手法では、バイラテラル重みと視点間の幾何学的な整合性を用いてマッチングスコアを高信頼化する。そして、3 つ目の手法では、出力されたパラメータに対してフィルタリングを行うことで外れ点や欠損の発生を抑制する。公開データセットを用いた精度評価実験により、従来の MVS 手法と比較を行うとともに、提案手法が撮影環境や復元対象の種類を問わずに高精度な 3 次元復元を行えることを示す。

2. PatchMatch 多視点ステレオによる 3 次元復元

PM-MVS は、入力の多視点画像のそれぞれに対するデプスマップと法線マップを推定することで 3 次元復元を行う手法である。PM-MVS では、はじめに入力画像の各画素に対してデプスと法線を一樣乱数で与えた後、高いマッチングスコアを持つデプスと法線の伝播および乱数によるデプスと法線の調整を繰り返し行うことで、デプスマップと法線マップの高速かつ高精度な推定を可能にしている。

3. PM-MVS の高精度化アルゴリズム

本論文では、PM-MVS の高精度化手法として、(i) マッチングスコアと視点間の幾何学的関係およびメッシュに基づく画素単位での視点選択手法、(ii) バイラテラル重み付き NCC (Bilateral-weighted NCC: BNCC) と Geometric Consistency による高信頼なマッチングスコアの算出手法、(iii) 重み付きメディアンフィルタと多視点幾何の整合性に基づくフィルタを用いたフィルタリング手法の 3 つを提案する。(i) では、視点選択のロバスト性を向上させるために、マッチングに用いる視点を画素単位で選択する。視点を選択する際には、マッチングスコア、視点間の輻輳角、物体面の法線およびメッシュを考慮し、隣接する視点の中から最適な視点を選択する。(ii) では、物体境界や低テクスチャ領域における誤対応点の発生を抑制するために、BNCC と Geometric Consistency を用いてマッチングスコアを計算する。(iii) では、デプス

表1 データセット terrace の実験結果：各手法の閾値 tol. [mm] における F_1 -score (F_1) [%]

tol.	PMVS F_1	COLMAP F_1	Yodokawa F_1	提案手法 F_1
1	9.81	47.77	58.10	62.47
2	36.13	73.96	75.86	79.23
5	69.35	87.81	88.42	91.17
10	79.12	92.58	93.47	95.94
20	84.10	96.33	96.88	98.15
50	88.40	99.38	98.57	99.08



図1 データセット terrace における (a) 入力画像例, (b) 真値の点群

マップ, 法線マップと点群に対するフィルタリングの処理として, 重み付きメディアンフィルタと多視点幾何の整合性に基づくフィルタを追加する. フィルタリングによってデプスマップや法線マップに含まれる外れ値を除去し, 3次元点群に含まれる外れ点を除去することで, 復元精度のさらなる向上を図る.

4. 性能評価実験

実験には, ETH3D データセット³⁾の“terrace”を用いる. terrace は, 屋外のテラスを撮影したデータセットであり, 23 視点から撮影された画像である. また, レーザースキャナで計測された真値の点群データが提供されている. 入力画像例と真値の点群を図1に示す. 本実験では, SfM ツールである COLMAP⁴⁾により推定された各視点のカメラパラメータを入力として用いる. 表1に F_1 -score (F_1) の評価結果を, 図2に各手法の復元結果を示す. 表1より, 提案手法は多くの閾値 tol. において F_1 -score が最も高い結果になっていることがわかる. また, 図2より, 提案手法の復元結果には外れ点が従来手法より多く含まれるが, 復元される領域が広く, 密な点群が復元されている.

次に, 復元が困難な物体である耳の3次元復元に提案手法を応用する. 実験に用いた画像の例と復元結果を図3に示す. 図3より, 復元結果にはオクルージョンの影響による欠損が見られるものの, 大部分の形状が復元できていることが確認できる.

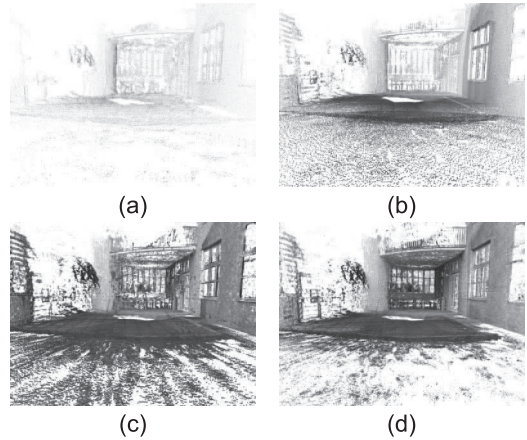


図2 データセット terrace における各手法の3次元復元結果：(a) PMVS, (b) COLMAP, (c) 淀川らの手法, (d) 提案手法

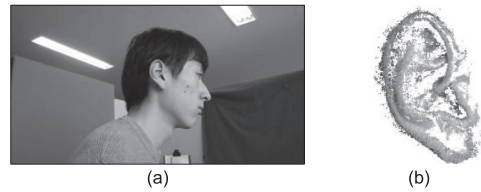


図3 耳の3次元復元：(a) 入力画像例, (b) 復元結果

5. まとめ

本論文では, PatchMatch 多視点ステレオの3つの高精度化手法を提案し, 公開データセットを用いた性能評価実験を通して, 提案手法の有効性を実証した. 今後の展望として, 芸術分野での3次元モデル制作への応用が考えられる.

文献

- 1) M. Hiradate, K. Ito, T. Aoki, T. Watanabe, and H. Unten, “An extension of PatchMatch Stereo for 3D reconstruction from multi-view images,” Proc. Asian Conf. Pattern Recognition, pp.061–065, Nov. 2015.
- 2) 伊藤隆史, 伊藤康一, 青木孝文, “低テクスチャ領域を含む多視点画像からの高精度デプス推定に関する検討,” 第23回画像の認識・理解シンポジウム, no.IS2-1-5, Aug. 2020.
- 3) T. Schöps, J.L. Schönberger, S. Galliani, T. Sattler, K. Schindler, M. Pollefeys, and A. Geiger, “A multi-view stereo benchmark with high-resolution images and multi-camera videos,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp.2538–2547, July 2017.
- 4) J.L. Schönberger and J.-M. Frahm, “Structure-from-motion revisited,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp.4104–4113, June 2016.